

University of Groningen

## Density functional theory applied to copper proteins

Swart, Marcel

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2002

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Swart, M. (2002). *Density functional theory applied to copper proteins*. [Thesis fully internal (DIV), Groningen]. s.n.

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

*chapter six*

# COPPER PROTEINS

*azurin*

## CONTENTS

6.1	<i>Redox chemistry</i>	<i>115-123</i>
6.2	<i>Force constants</i>	<i>124-127</i>
6.3	<i>Axial bonding</i>	<i>128-131</i>
6.4	<i>Vibrational frequencies</i>	<i>132-139</i>

## SUMMARY

*The creation of a copper force field for use in azurin (and derivatives) is described in the first two sections of Chapter 6, and is followed by sections concerning the axial bonding in the active site of azurin and the validation of the copper force field in Molecular Dynamics of azurin. The creation of the copper force field involves finding suitable atomic charges for the reduced and oxidized state and at stages in-between these states. The effect of the electric field of the charges in the surrounding protein on the energy and atomic charges is checked, and the force constants for copper-ligand interactions determined. In the last section, the vibrational frequencies from either DFT, the copper-ligand force constants or the ones obtained in MD simulations are compared.*

## 6.1

**Redox chemistry**

*Charge distribution and energy changes of active sites of copper proteins in redox reactions*

In this chapter several copper proteins are investigated by a combined Density Functional Theory (DFT)<sup>1</sup> / Molecular Dynamics (MD)<sup>160</sup> approach. The force field parameters needed for describing the active site in the classical MD simulations are obtained from DFT calculations on the active site, either with or without the rest of the protein present. Density Functional Theory calculations are used since it is an accurate yet computationally cheap method, that enables a treatment of larger systems (in this case the complete active site) at a high computational level. Moreover, as the basic property of DFT is the charge density, it enables to perform calculations on the active site where an electron is removed in steps of 0.2 electron in going from the reduced to the oxidized state. This property of DFT is impossible (or at least very difficult and awkward) to obtain with wavefunction (Hartree-Fock) based methods.

The force field parameters are thus obtained as a function of the removal of the electron, and can be used in MD simulations where an electron is removed in steps from the system. Using thermodynamic integration it is then possible to obtain the redox potential of the copper protein, i.e. the change in free energy in going from the reduced to the oxidized state. This change in free energy  $A$  should be obtained after integration over *all* degrees of freedom of the system in a Molecular Dynamics simulation<sup>160</sup>, in which the partition function  $Q_{NVT}$  is obtained:

$$A = -kT \ln Q_{NVT}$$

$$Q_{NVT} = \int \exp \left( -\frac{U_i}{kT} \right) d\mathbf{r}_i \quad (1)$$

$$\frac{A}{\lambda} = \frac{-kT}{Q_{NVT}} \frac{Q_{NVT}}{\lambda} = \frac{-kT}{Q_{NVT}} \frac{-1}{kT} \int \frac{U_i}{\lambda} e^{-U_i/kT} d\mathbf{r}_i = \left\langle \frac{U_i}{\lambda} \right\rangle_{ens}$$

Integrating the derivative of the Helmholtz free energy  $A$  with respect to  $\lambda$  from zero (reduced state) to one (oxidized state), the redox potential is obtained. Experimentally, the redox potential is measured relative to hydrogen, which makes it difficult to compare the absolute values of the redox potentials directly. However, it is possible to compare computed and experimentally found redox potential *differences*.

The copper protein studied is azurin. Although its function in nature is still unclear, it presumably serves as an electron transfer protein. Its structure was solved only in the last twenty-five years<sup>6,7,277,278</sup>, and showed three in-plane ligands and two axial bonding groups for the copper ion. The in-plane ligands are two histidines and one cysteine (see Figure 6.1.1). Above the plane resides a methionine; below the plane the backbone oxygen of a glycine is oriented towards the copper. The first four of these residues seem to be common for the type 1 copper proteins, while the fifth group (glycine) is found only in azurin. In Section 6.3, the nature of the interactions between the copper and the axial ligands will be discussed in more detail.

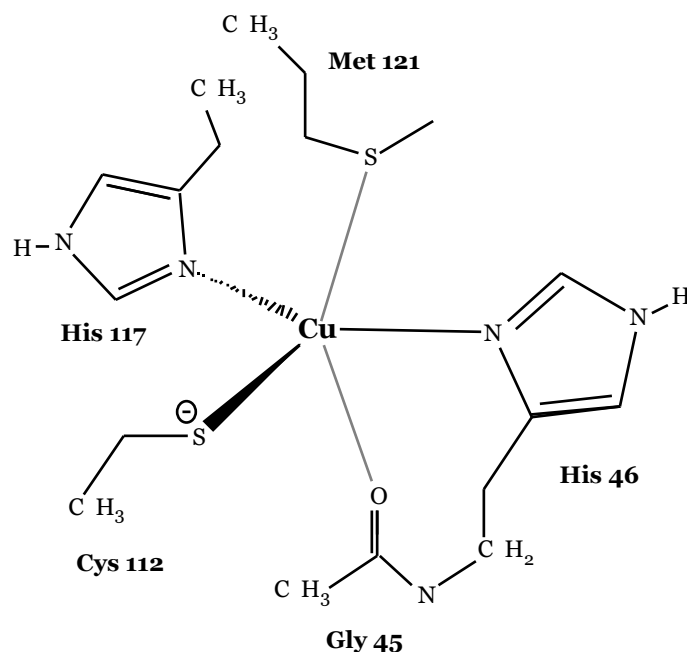


FIGURE 6.1.1. ACTIVE SITE OF AZURIN

In the past, site-directed mutagenesis has been used to study the influence of amino acid residues on the stability and properties of the active site, by specifically replacing residues, both in- and outside the active site. These studies showed a marked dependence of the reduction potential on the type of residue at the 121 position, in a range of 210 to 520 mV<sup>3</sup>.

The azurin molecules studied in this section are from *Pseudomonas aeruginosa* (*Pa*) and are *wildtype* at low and high pH (5.5 and 9.0 respectively)<sup>7</sup>, the N47D (Asn → Asp) mutant<sup>4</sup> and the F114A (Phe → Ala) mutant<sup>5</sup>. The pH-dependence of *wildtype Pa* azurin has been studied thoroughly<sup>7,72</sup>, showing a deprotonation at high pH of the His35 residue, which is located at some 12 Å from the copper ion. This deprotonation seems to have only a negligible effect on the active site geometry<sup>7,72</sup> but does involve a remarkable backbone shift around the His35 position (Pro36-Gly37 peptide flip)<sup>7</sup>. In the calculations in this section the pH can not be taken explicitly into account; it can only be mimicked by using the protein coordinates from the X-ray crystallographic data at either high or low pH, and use an appropriate protonation state for the His35 residue. Although the active site in both cases is the same, the influence of the surrounding protein should be different. The two *wildtype* molecules will be referred to as *wt1* (active site and protein coordinates taken from X-ray data at low pH) and *wt2* (coordinates taken from X-ray data at high pH) from this point on.

The N47D-mutant PDB-file contains a copper ion, although the structure itself was obtained for the zinc derivative, which resembles that of zinc-substituted *wildtype* azurin<sup>59</sup>. Optical and EPR spectra for the copper derivative of the N47D-mutant are similar to *wildtype* and, under the assumption that these should change considerably if the copper ion moves around, this seems to indicate that the structure of the copper derivative is similar to the *wildtype* structure.

## Computational details

The Density Functional Theory calculations were performed with the ADF program (versions 2.3.3, 1999.03, 2000.02)<sup>177,178,187</sup> on a cluster of either IBM RS/6000 or Pentium-Linux boxes. In all calculations the Becke<sup>120</sup>-Perdew<sup>121</sup> exchange-correlation potential was used in a triple zeta basis set with polarization functions (TZP). To study the charge distribution in the active site, the Multipole Derived Charge analysis<sup>183</sup> (see Section 3.1) was used, where the MDC-q charges have been used.

The coordinates for the four azurin variants were taken from PDB-files that were obtained from the Brookhaven Protein Database {F114A *Pa* (1azn), N47D *Pa* (1azr), *wildtype Pa* obtained at pH5.5 (4azu; hereafter referred to as *wt1*), *wildtype Pa* obtained at pH9.0 (5azu; *wt2*)}

## Results

The coordinates for the atoms in the active site were taken from the PDB-files, with the amino acid residues cut off at the carbon-alpha position. Hydrogen atoms were added where appropriate and the hydrogen positions optimized while the other positions were kept fixed. This ensured that all experimental data stayed as they were, while only the unknown positions were added. The other positions could have been optimized in our calculations also, but as the rest of protein was not included, the residues would not feel the interactions with the rest of the protein. In Section 6.3 and Chapter 9, this issue will be addressed in more detail.

In all PDB-files, four molecules were found in the asymmetric unit, leading to four subunits, and therefore to four active site structures per azurin variant. Each of these subunits has been treated separately and the results averaged afterwards.

### Energies

After optimization of the hydrogen positions, the energies were computed when an electron was removed in steps of 0.2 electron from each active site structure. The energy profile as a function of the removal of the electron is given in Figure 6.1.2 for the four azurin variants. In these plots, the energies are given as a function of the total charge of the system ( ), which corresponds to 0.0 for the reduced state and +1.0 for the oxidized state.

The removal of the electron is accompanied by an increase in energy that fits perfectly to a second degree polynomial. For *wt2* and N47D azurin, a small r.m.s. variation is found for the energies of the four subunits of respectively 1.36 and 1.35 kcal/mol. For *wt1* and F114A azurin, these variations are a bit larger (6.32 and 7.40 respectively). The energy needed for going from the reduced to the oxidized state (ionization potential), is rather similar for the four azurin variants (see *in vacuo* results in first column of Table 6.1.1). The difference between the *wt1* and *wt2* energies is negligible, while the difference with the mutants is still rather small (less than 3 kcal/mol). For comparison, also the ionization potential of an isolated Cu ion is given in Table 6.1.1, which has a much higher value of almost 500 kcal/mol.

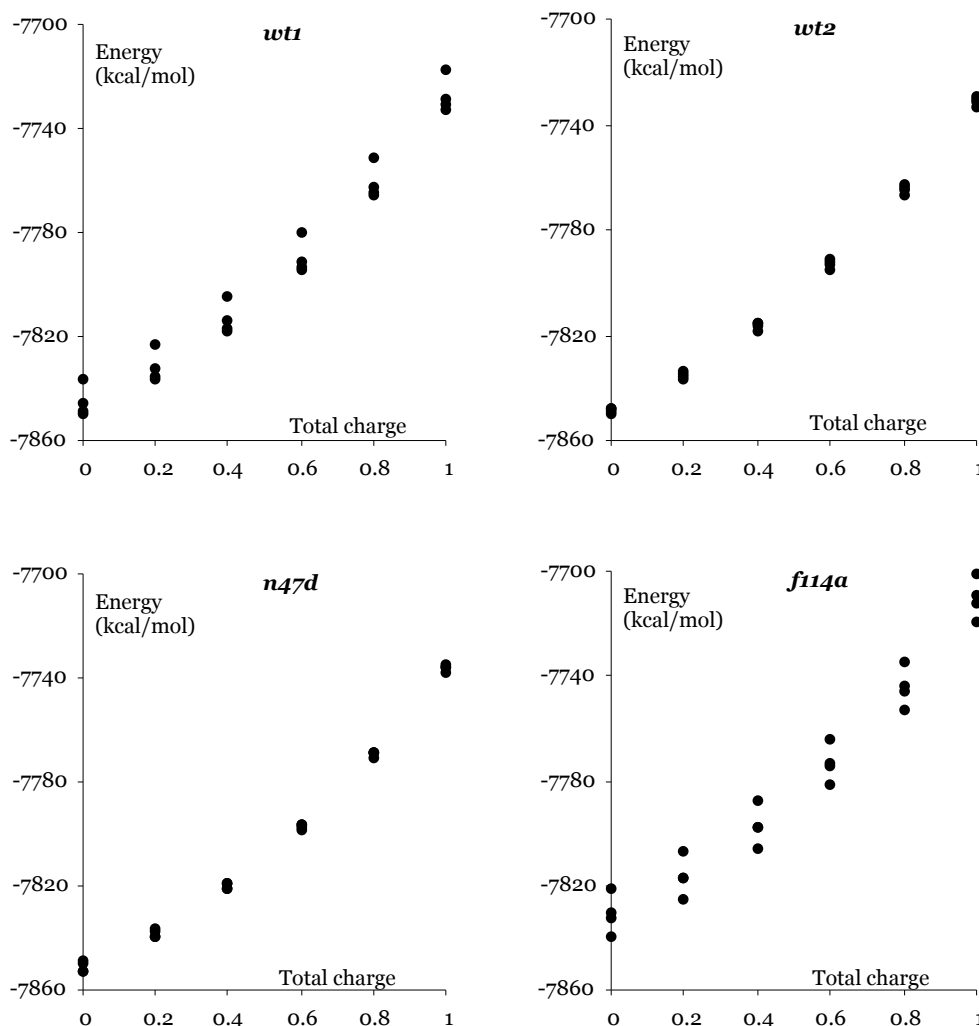


FIGURE 6.1.2. ENERGY PROFILE<sup>a</sup> FOR REMOVAL OF ELECTRON FROM ACTIVE SITE *IN VACUO*

The rest of the protein was taken partly into account in subsequent calculations, in which it was represented by point charges only. The interaction between the active site and the rest of the protein was therefore reduced to electrostatics only. However, as the coordinates of the atoms in the rest of the protein were kept fixed, the only term in the interaction energy (in non-polarizable force fields like AMBER<sup>126</sup> or GROMOS<sup>127</sup>) that changes in going from the reduced to the oxidized state, is the electrostatic energy. Two sets of protein charges were used, either the standard Gromacs<sup>279</sup> charges or the charges obtained with the Multipole Derived Charge analysis<sup>183</sup> for isolated amino acid residues.

The inclusion of the point charges in the calculations has a stabilizing effect on the active sites of all subunits. The energy goes down for all for them, but not by the same amount for all azurin variants and not by the same amount for the reduced or oxidized state. Both result from the total charge contained in the rest of the protein. This is either  $-2$  or  $-3$ , depending on the variant and the pH at which the crystal structure was obtained. For *wt1* and N47D azurin it is  $-2$ , while it is  $-3$  for *wt2* azurin. F114A has two subunits with a

<sup>a</sup> The four dots at every point correspond to the four molecules in the asymmetric unit

protein charge of  $-2$  and two with  $-3$ . Due to the negative protein charge, the oxidized state (with a positive charge extra) is stabilized more than the reduced state. This leads to a decrease of the ionization potential (see Table 6.1.1).

TABLE 6.1.1. COMPUTED IONIZATION ENERGIES (KCAL/MOL)

	<i>in vacuo</i>	<i>GMX surroundings</i>	<i>MDC surroundings</i>
<i>wt1</i>	117.63 $\pm$ 0.77	92.88 $\pm$ 4.83	71.64 $\pm$ 4.79
<i>wt2</i>	117.62 $\pm$ 0.76	55.82 $\pm$ 4.70	36.25 $\pm$ 4.35
n47d	114.83 $\pm$ 2.26	48.04 $\pm$ 1.67	48.04 $\pm$ 0.72
f114a	120.47 $\pm$ 2.01	91.68 $\pm$ 9.30 <sup>*</sup>	67.24 $\pm$ 8.15 <sup>*</sup>
		47.63 $\pm$ 0.69 <sup>#</sup>	25.45 $\pm$ 2.52 <sup>#</sup>
Cu <sup>+</sup>	499.33		

GMX) Gromacs    \*) Protein charge  $-2$     #) Protein charge  $-3$

### Charge distribution

For the active sites *in vacuo* (without the protein environment charges present), the charge distribution has been calculated using the Multipole Derived Charge analysis. Just like the calculated energy, this is also obtained as a function of the removal of an electron from the site. Given in Table 6.1.2 are the residue charges, i.e. the total charge contained in a certain amino acid residue; the distribution of charge within the residues is depicted in Figure 6.1.3 for the reduced state and in Figure 6.1.4 for the oxidized state.

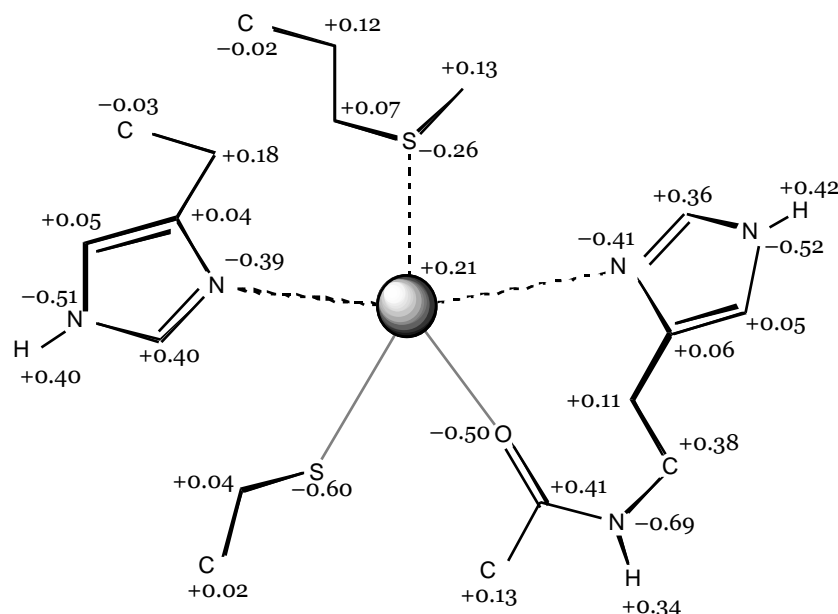


FIGURE 6.1.3. CHARGE DISTRIBUTION IN ACTIVE SITE (REDUCED STATE)



The difference between the four azurin variants is rather small, just like the energies of the active site *in vacuo*. The r.m.s. variation of the residue charges for all 16 active site structures (four subunits of four azurin variants) is rather small with values of  $\sim 0.02$ - $0.4$  a.u.

TABLE 6.1.2. MDC-Q CHARGES FOR RESIDUES IN ACTIVE SITE

	<i>vacuo</i>	<i>GMX surroundings</i>	<i>MDC surroundings</i>
<i>reduced state</i>			
Gly45	$0.041 \pm 0.025$	$0.030 \pm 0.026$	$0.010 \pm 0.025$
His46	$0.101 \pm 0.022$	$0.166 \pm 0.077$	$0.200 \pm 0.059$
Cys112	$-0.541 \pm 0.040$	$-0.598 \pm 0.074$	$-0.566 \pm 0.067$
His117	$0.141 \pm 0.024$	$0.151 \pm 0.032$	$0.170 \pm 0.039$
Met121	$0.049 \pm 0.020$	$0.019 \pm 0.022$	$-0.023 \pm 0.026$
Cu	$0.209 \pm 0.042$	$0.233 \pm 0.040$	$0.209 \pm 0.036$
<i>oxidized state</i>			
Gly45	$0.060 \pm 0.028$	$0.068 \pm 0.025$	$0.033 \pm 0.026$
His46	$0.239 \pm 0.020$	$0.311 \pm 0.074$	$0.337 \pm 0.057$
Cys112	$-0.030 \pm 0.047$	$-0.088 \pm 0.088$	$-0.047 \pm 0.076$
His117	$0.268 \pm 0.017$	$0.281 \pm 0.028$	$0.307 \pm 0.034$
Met121	$0.135 \pm 0.049$	$0.095 \pm 0.028$	$0.053 \pm 0.026$
Cu	$0.329 \pm 0.043$	$0.334 \pm 0.036$	$0.317 \pm 0.034$
<i>ox - red</i>			
Gly45	0.019	0.038	0.024
His46	0.138	0.145	0.137
Cys112	0.511	0.510	0.518
His117	0.126	0.130	0.137
Met121	0.086	0.076	0.075
Cu	0.120	0.101	0.108

GMX) Gromacs

In the reduced state, the total charge on the axial groups is small ( $\sim 0.05$  a.u.), and a bit larger on the histidines ( $0.10$ - $0.14$  a.u.); Cys112 has a rather large total negative charge of  $-0.541$  a.u. The copper ion is found not to have its formal charge in either the reduced or oxidized state; instead, its charge is  $0.21$  in the reduced state and  $0.33$  in the oxidized state due to back-donation of charge from the ligands. The largest contributor to the back-donation is found to be Cys112 that accounts for more than half of the electron being removed. The other two in-plane ligands account for some 25 % more, while the remainder is coming mostly from Met121 and copper itself; Gly45 accounts only for 2 %. This leads to a total charge on Cys112 around zero, while both histidines have a charge of around  $0.24$ - $0.27$  a.u. The charge on Gly45 remains small, while some considerable amount of positive charge is found on Met121.

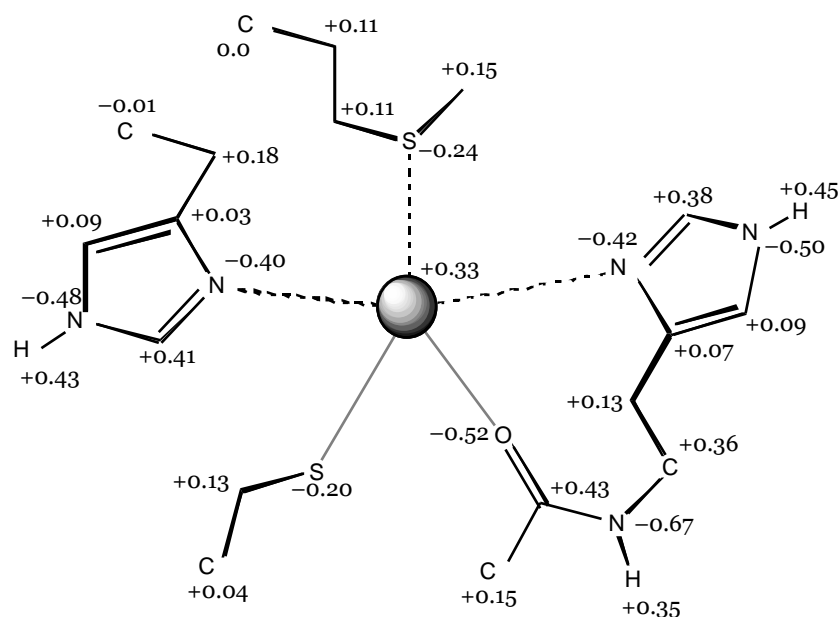


FIGURE 6.1.4. CHARGE DISTRIBUTION IN ACTIVE SITE (OXIDIZED STATE)

The inclusion of the protein charges does not have such a marked effect on the charge distribution. Although there are some subtle changes of a few hundreds a.u., globally the picture remains the same as in the *vacuo* case, although the standard deviations are a bit higher now. However, the redox difference, i.e. the difference between the reduced and oxidized state, remains virtually the same for all residues.

## Discussion

The completely static inclusion of the protein environment can not account for the experimentally observed redox potential differences upon pH changes or site-directed mutations, as the protein is not allowed to relax in this section. For instance, the change in redox state is likely to induce changes in the protein-site interactions, e.g. through reorientation of polar groups. The solvent, which may also have a pronounced effect, has also not been taken into account in this section. However, it can be established if the presence of the protein environment has an effect on the active site energy or the charge distribution in it, or on both.

## Energies

A direct connection between the total protein charge and the stabilization by the protein charges can be made. The ionization potentials for *wt2* azurin and the two subunits of F114A azurin with a protein charge of  $-3$  are similar, as are the values for *wt1* azurin and the two subunits of F114A azurin with a protein charge of  $-2$ . At first sight, the values for N47D look odd, as they seem to indicate a total protein charge of  $-3$ , while in fact a value of  $-2$  is found. But the N47D mutant is in fact a double mutant, with a second mutation at a position far away from the active site (Glu2  $\rightarrow$  Gln). Therefore, the protein charge of the close by residues is indeed  $-3$ , as anticipated from the ionization potential change, while the total protein charge is the sum of this close by protein charge and the charge of residue 2.

The inclusion of the point charges leads to a stabilization of both the reduced and oxidized state, where the oxidized state is stabilized more than the reduced state (see Figure 6.1.5). Taking the reduced state *embedded* in the protein charges as reference state (*embed.red* in Figure 6.1.5), the additional stabilization when an electron is being removed in going to the *embedded* oxidized state (along path **C**), can be modeled quite accurately (average absolute deviation 0.7 kcal/mol along path **C**) by classical electrostatic interactions between the active site and the protein charges, when the MDC-q charges are used for the active site atoms.

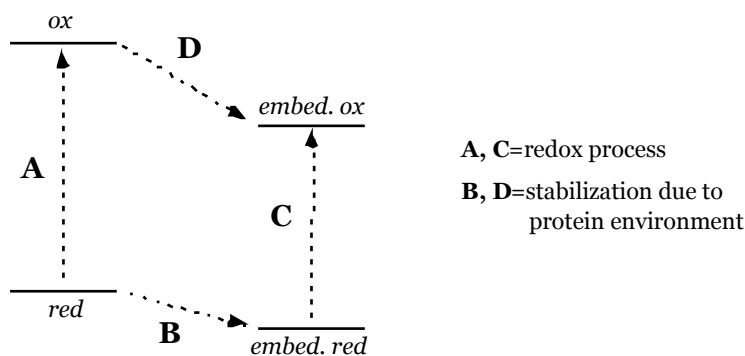


FIGURE 6.1.5. REDOX ENERGY PROFILE FOR EMBEDDING IN PROTEIN ENVIRONMENT

The stabilization of the reduced state due to the presence of the protein environment (path **B**) however can not be modeled by electrostatic interactions between the active site and protein charges alone; a subunit dependent *extra* stabilization of some 6-21 kcal/mol remains. This *extra* stabilization is most likely the result of polarization effects, which have not been taken into account in the classical electrostatics modeling.

The inclusion of protein charges therefore leads consequently to a lower energy, but the amount of stabilization depends on the azurin variant as well as on the (reduced or oxidized) state of the active site. This is just an energetic effect, which can be modeled quite well by classical electrostatics; the latter enables the use of classical MD simulations for obtaining redox potentials of copper proteins, when appropriate (MDC-q) charges for the active site are used.

#### Charge distributions

The r.m.s. variation of the (MDC-q) residue charges over the four azurin variants is rather small, especially if the differences in the active site geometries are taken into account. The distance of Gly45 to copper for instance varies from 2.3 to 3.1 Å; for Met121 it is in the range of 2.8-3.5 Å. Not only these weakly bonded axial groups show a considerable flexibility, also the distances of the in-plane ligands to copper vary to some extent. For His46, it is in the range from 2.0 to 2.2 Å, for Cys112 from 2.2-2.3 Å and for His117 from 1.9 to 2.6 Å. Still, the deviations in the charge distributions are rather small at 0.02-0.04 a.u.

The influence of including the protein charges in the DFT calculation is small; the changes in residue charges are at most some 0.1 a.u. However, again this change in residue charge is almost constant for all four azurin variants. The standard deviation increases just by a small amount. There is also hardly any difference between the two sets of protein charges; the residue charges after inclusion of either the Gromacs or MDC charges are rather similar. The difference between the reduced and oxidized state charges (redox difference) is

even more similar for the three environments (*in vacuo*, *Gromacs* or *MDC* protein charges). The difference between the three environments is at maximum some 0.02 a.u. This is a very promising result, as it means that in the MD simulations, the standard *Gromacs* charges can be used for the protein, while the *MDC-q* charges should be used for the active site.

The inclusion of protein charges does not have a marked effect on the charge distribution in the active site. There are some small differences, but they are the same for the four azurin variants.

## Conclusions

The active site of the copper protein azurin has been studied with Density Functional Theory calculations for four variants: *wildtype* (*wt1* and *wt2*) and two mutants (N47D and F114A). The energy and charge distribution was computed along the process of removal of an electron from the system in steps of 0.2 electron. These calculations were performed either *in vacuo* or surrounded by the charges of the rest of the protein, for which two sets of charges were used, the standard *Gromacs* charges or charges obtained for amino acid residues from the Multipole Derived Charge analysis.

The ionization potentials of the active sites *in vacuo* are rather similar, with a deviation of at most 3 kcal/mol for the mutants. The ionization potential for *wildtype* azurin has been obtained for active site structures that have been determined experimentally at low (5.5) and high pH (9.0). This pH difference was found experimentally to cause a deprotonation of the His35 residue, as well as to have a negligible effect on the active site. It is confirmed by the DFT results, which give an ionization potential of 117.6 kcal/mol in both cases for the active site *in vacuo*. Because of the classical electrostatic effect, differences are observed when protein charges are included in the quantum calculation.

The inclusion of protein charges has a stabilizing effect on the active site, the energy is lowered by a subunit and state dependent amount. These stabilizations can be modeled quite accurately by classical electrostatics. The protein charges have a negligible effect on the charge distribution in the active site. This feature can be used in MD simulations, where the standard *Gromacs* charges can be used for the protein charges, while the *MDC-q* should be used for the atoms in the active site.

## Force constants

Force field parameters for copper for several azurin molecules

In Section 6.1, Density Functional Theory<sup>1</sup> (DFT) was used to find the energy and charge distribution for the active site of a few azurin molecules. The resulting charges can be used directly in classical Molecular Dynamics (MD) simulations<sup>160</sup> for a proper description of the non-bonded interactions (together with standard Lennard-Jones parameters<sup>127</sup>). However, for the bonded interactions there are no standard force field parameters available, since the interactions of metals are difficult to generalize in terms of a simplified force field. Therefore, the force field parameters for describing the bonded interactions of copper with its surrounding residues should be obtained from DFT calculations.

As a first attempt, a simple approach has been tried, in which the DFT Hessian of the Cu atom alone (a  $3 \times 3$  matrix with six independent values) was used to fit force constant values of copper-residue bonds. Five bonds connecting copper to the residues in the active site have been used, using a harmonic bond potential for each of them. The second derivative of these potentials with respect to the atomic coordinates provides a (force field, FF) Hessian matrix, which is a function of the force constants and the atomic coordinates (see Appendix of Section 3.2 for explicit formulas). The force constants for the five bonds were fitted to minimize the difference between the DFT and FF Cu-Hessians<sup>a</sup>. Although this procedure seems to give a reasonable Cu-Hessian for one *wildtype* structure, the obtained force constants look odd with a force constant for the “weakly bonded” Gly45 (at 2.96 Å) of 44 kcal/mol/Å<sup>2</sup>, while for the in-plane ligand His46 (at 2.06 Å) a value of only 31 kcal/mol/Å<sup>2</sup> is found. Using this procedure for other active site structures even results in negative (!) force constants for some of the residues. Therefore, fitting the Cu-Hessian is not providing reliable force constants. Subsequently, a new method (IntraFF) has been developed, described in more detail in Section 3.2, which does take the bonding interactions into account on *both* sides, and provides reliable force constants. In the IntraFF method, only those parts of the DFT Hessian are needed that are related to the atoms involved in the bonds; as we want to have force constants for bonds of copper to the nearest atom in the residues (O-Gly45, N-His46, S-Cys112, N-His117, S-Met121), only the parts of the DFT Hessian related to these six atoms are needed (therefore a  $18 \times 18$  matrix).<sup>b</sup>

### Computational details

The Density Functional Theory calculations were performed with the ADF program (versions 2.3.3, 1999.03, 2000.02)<sup>177,178,187</sup> on a cluster of either IBM RS/6000 or Pentium-Linux boxes. In all calculations the Becke<sup>120</sup>-Perdew<sup>121</sup> exchange-correlation potential was used in a triple

---

<sup>a</sup> In principle, the DFT Hessian should be corrected for non-bonding (electrostatic and Lennard-Jones) interactions between copper and the residue atoms with which copper does not have bonding interactions; however, these corrections are of the order of  $10^{-4}$  a.u. for the Hessian, while the DFT Hessian values for copper are of the order of  $10^{-1}$  a.u.

<sup>b</sup> Of course one needs the complete active site residues in the DFT calculations; therefore, although the Hessian is obtained only for the Cu-O(Gly45)-N(His46)-S(Cys112)-N(His117)-S(Met121) atoms, it is obtained for a system in which all active site residue atoms are present

zeta basis set with polarization functions (TZP). As the Hessian is calculated in the ADF program<sup>117,187</sup> through numerical differentiation of the analytical gradients, obtaining the Hessian matrix takes a long time. Therefore, for each azurin variant one subunit was selected for which the Hessian was computed. Afterwards, the IntraFF analysis was performed using a standalone program.

In this section, the molecules used in Section 6.1 are considered, and some other azurin variants have been added. The additional molecules are the M121Q (Met121 Gln)<sup>42</sup> and N47L<sup>84</sup> (Asn47 Leu) mutants. The coordinates for the azurin variants were taken from PDB-files that were obtained from the Brookhaven Protein Database {F114A *Pseudomonas aeruginosa* (1azn), N47D *Pseudomonas aeruginosa* (1azr), wildtype *Pseudomonas aeruginosa* pH5.5 (4azu; hereafter referred to as *wt1*), wildtype *Pseudomonas aeruginosa* pH9.0 (5azu; *wt2*)} or from personal communication<sup>280</sup> {*Alcaligenes denitrificans* M121Q Ad, N47L Ad}.

## Results

The IntraFF harmonic force constants for the azurin variants in both the reduced and oxidized state, with corresponding “equilibrium” bond distances are given in Table 6.2.1. The “equilibrium” bond distances were simply taken as the values as they are obtained in the X-ray data (which were obtained in the oxidized state).

TABLE 6.2.1. INTRAFF FORCE CONSTANTS  $K$  (KCAL/MOL/Å<sup>2</sup>) AND DISTANCES  $R$  (Å)

	<i>Gly45</i>	<i>His46</i>	<i>Cys112</i>	<i>His117</i>	<i>Met/Gln-121</i>
<i>wt1</i>					
R(Cu-L)	2.955	2.064	2.267	1.978	3.164
K(red.)	26.3	67.2	148.9	124.1	30.6
K(ox.)	30.5	115.9	168.9	177.7	36.9
<i>wt2</i>					
R(Cu-L)	2.828	1.976	2.284	2.056	3.147
K(red.)	24.0	128.4	146.7	68.8	31.9
K(ox.)	28.0	181.7	166.9	117.8	39.3
<i>n47d</i>					
R(Cu-L)	2.362	2.087	2.218	1.869	3.440
K(red.)	21.2	61.3	199.6	246.6	20.8
K(ox.)	32.9	109.4	223.5	304.6	33.3
<i>f114a</i>					
R(Cu-L)	3.066	2.208	2.216	2.362	3.007
K(red.)	14.1	72.5	179.3	16.2	36.1
K(ox.)	16.5	89.3	179.0	17.5	39.1
<i>n47l</i>					
R(Cu-L)	2.951	1.975	2.130	1.943	3.121
K(red.)	39.9	164.5	290.4	169.1	32.7
K(ox.)	49.2	226.5	323.8	229.5	45.3
<i>m121q</i>					
R(Cu-L)	3.468	1.907	2.134	2.064	2.250
K(red.)	40.3	243.1	298.2	82.7	37.7
K(ox.)	59.2	300.8	324.3	124.7	55.3

## Discussion

The IntraFF force constants for the copper to residue bonds are consistently larger in the oxidized state than in the reduced state, indicating stronger bonding in the former. Furthermore, the force constants for the in-plane ligands are always larger than for the axial groups, except for His117 in the F114A mutant. However, in that case the Cu-ligands is elongated to a large extent at 2.36 Å.

The Cu-S(Cys112) force constant of  $\sim 170$  kcal/mol/Å<sup>2</sup> as well as the Cu-N force constants of 115–180 kcal/mol/Å<sup>2</sup> for *wildtype* azurin are small compared to “normal” bonds. For instance, the force constant for a C-C single bond is 620 kcal/mol/Å<sup>2</sup> and for a double bond even 938 kcal/mol/Å<sup>2</sup> in the AMBER95<sup>126</sup> force field, while a C-S bond has a force constant of 474 kcal/mol/Å<sup>2</sup>. Still, the Cu-Cys/His force constants are considerably larger than the Cu-Gly45/Met121 bonds (30–40 kcal/mol/Å<sup>2</sup>). The latter values are still too large to be represented by non-bonded interactions only in force field calculations; non-bonded interactions between copper and the active site residues result in Hessian values of  $1\text{--}5\cdot 10^{-4}$  a.u. while force constant values of  $\sim 30$  kcal/mol result in Hessian values  $1\cdot 10^{-2}$  a.u.

For the in-plane ligands (His46, Cys112 and His117) an interesting trend can be observed: the force constant increases if the “equilibrium” distance decreases, which is indicative for an anharmonic potential (for a pure harmonic potential the force constant would remain constant). For instance, in Figure 6.2.1 the force constants for His117 are plotted against the “equilibrium” distance.

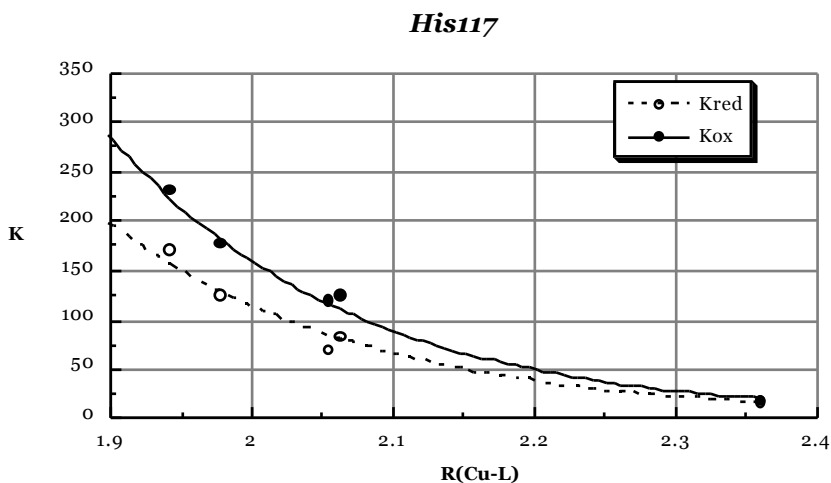


FIGURE 6.2.1. HIS117 INTRAFF FORCE CONSTANTS AS A FUNCTION OF THE DISTANCE

Fitted through the points on the plot is an exponential function, which fits nicely with the obtained points. A similar exponential function would have been found for the Hessian (in 1D) of either the Morse<sup>130</sup> or Frost<sup>131,132</sup> potential (see Section 3.2), which are both anharmonic bond potentials. Using the IntraFF method, Frost parameters were obtained for the five bonds in the four azurin variants; the average values of these parameters are given in Table 6.2.2 for both the reduced and oxidized state, together with the equilibrium distance and dissociation energy. Also given is the force constant for a harmonic potential at the Frost equilibrium distance, which is equal (in 1D) to the 2<sup>nd</sup> derivative of the Frost potential at the equilibrium distance.

TABLE 6.2.2. FROST PARAMETERS ( $Z_1, Z_2, a, b$ ; A.U.) FOR ACTIVE SITE OF AZURIN IN VACUO, WITH CORRESPONDING EQUILIBRIUM DISTANCE  $R_{EQ}$  (Å), DISSOCIATION ENERGY  $D_E$  (EV) AND FORCE CONSTANT  $K(R_{EQ})$  FOR HARMONIC POTENTIAL (KCAL/MOL/Å<sup>2</sup>)

	$Z_1$	$Z_2$	$a$	$b$	$R_{eq}$	$D_e$	$K(R_{eq})$
Gly45							
<i>reduced</i>	29	8	1.37	44.8	3.08	0.05	9.1
<i>oxidized</i>	29	8	1.17	51.4	2.78	0.42	62.4
His46							
<i>reduced</i>	29	7	1.60	65.3	1.93	0.76	214.9
<i>oxidized</i>	29	7	1.51	65.1	1.95	1.03	263.4
Cys112							
<i>reduced</i>	29	16	1.51	130.2	2.19	0.96	237.3
<i>oxidized</i>	29	16	1.44	133.7	2.15	1.50	346.1
His117							
<i>reduced</i>	29	7	1.62	65.3	1.93	0.70	203.7
<i>oxidized</i>	29	7	1.53	65.6	1.93	1.03	269.0
Met121							
<i>reduced</i>	29	16	1.10	94.0	3.03	0.63	83.9
<i>oxidized</i>	29	16	1.09	92.4	3.08	0.61	78.7

## Conclusions

Force field parameters for the bonding interactions of copper in a few azurin variants have been obtained from Density Functional Theory calculations. The force constants were obtained with the IntraFF method, which extracts them from a quantum chemically computed Hessian matrix. This method provides reliable force constants that can be used in classical Molecular Dynamics simulations.

The obtained force constants are consistently larger in the oxidized state than in the reduced state, indicating stronger bonding in the former. Furthermore, the force constants show anharmonic behavior, i.e. the force constant values increase with decreasing “equilibrium” distance. The parameters for the anharmonic Frost potential have been obtained with the IntraFF method for all azurin variants and averaged to obtain a generally applicable copper (azurin) force field.



## Axial bonding

*Checking the bonding character of the axial groups and monitoring the potential energy surface*

In this section, the bonding character of the axial groups is investigated by looking at the mixing of the copper and axial residue orbitals. The residue-copper distance of either the Met121 and the Gly45 residue has been systematically varied from 2.5 to 4.0 Å in steps of 0.1 Å, while the rest of the geometry was kept fixed. This enables a check on the influence of this distance on the mixing of copper-residue orbitals, and a potential energy surface is obtained for these two ligands, both in the reduced and the oxidized state.

### Computational details

The active site structure<sup>a</sup> used in this study was taken from the PDB-file of *wildtype Pa* azurin (B subunit) for which the X-ray data were obtained at pH 5.5<sup>7</sup>. The copper-residue distances for the residues in the active site are 2.064 Å (His46), 2.267 Å (Cys112), 1.978 Å (His117) for the in-plane ligands, and 2.955 Å (Gly45), 3.164 Å (Met121) for the axial ligands. The coordinates were transformed from Cartesian format into Z-matrix format<sup>1</sup>. This enables to vary either the Cu-O(Gly45) or the Cu-S(Met121) distance, and let the corresponding residue follow. In the following, either the Cu-Gly45 or the Cu-Met121 distance was varied, while the rest of the coordinates were kept fixed. As the Gly45 and His46 residues are connected, varying the Cu-Gly45 could have resulted in changing the Cu-His46 distance also (see Figure 6.3.1).

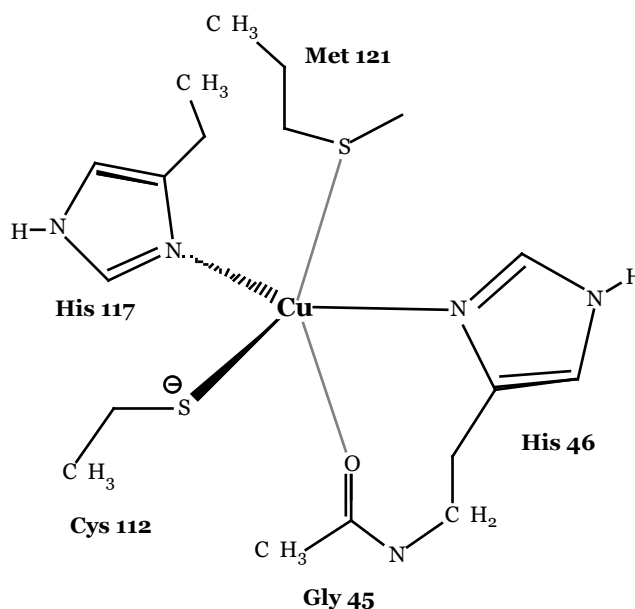


FIGURE 6.3.1. ACTIVE SITE OF AZURIN

<sup>a</sup> The active site consists of the same residues and atoms as described in section 6.1

In the used Z-matrix however, only the atoms of Gly45 changed position, while all atoms belonging to His46 remained at their position. The change in Cu-Gly45 distance was therefore accompanied by a change in the angles between the backbone of His46 and the atoms of Gly45.

The Density Functional Theory calculations were performed with the ADF program (version 2000.02)<sup>177,178,187</sup> on a cluster of Pentium-Linux boxes. In these calculations the Becke<sup>120</sup>-Perdew<sup>121</sup> exchange-correlation potential was used in a triple zeta basis set with polarization functions (TZP).

## Results and discussion

### *Variation of the Gly45-Cu distance*

For the reduced state with Gly45 at 2.5 Å, the highest occupied molecular orbital (HOMO), which is doubly occupied, consists for 51 % of contributions from the sulphur of Cys112 and 36 % from copper, and no contribution from the Gly45 oxygen. This oxygen has some contributions in orbitals where it is mixed with orbitals of its neighboring atoms, and a very small mixing (1 %) with Cu d-orbitals in one orbital (HOMO-2<sup>a</sup>). When Gly45 is placed at the other extreme of 4.0 Å, the HOMO has hardly changed: 54 % S<sub>112</sub> and 33 % Cu. Naturally, no mixing with copper orbitals are found at this distance.

For the oxidized state with Gly45 at 2.5 Å, the singly occupied HOMO consists again mainly of contributions from S<sub>112</sub> (52 %) and copper (28 %), and no contribution from O<sub>45</sub>. The oxygen atom however does have some small contributions in the HOMO-1 (1 %), and some more significant contributions in the HOMO-3 (14 %) that consists for 53 % of Cu d-orbitals, and the HOMO-10: 50 % O<sub>45</sub> and 18 % Cu d-orbitals. With the Gly45 at 4.0 Å, the HOMO consists mainly of copper d-orbitals (26 %) and S<sub>112</sub> orbitals (52 %). At this distance, the oxygen of Gly45 does of course not mix with the copper orbitals anymore.

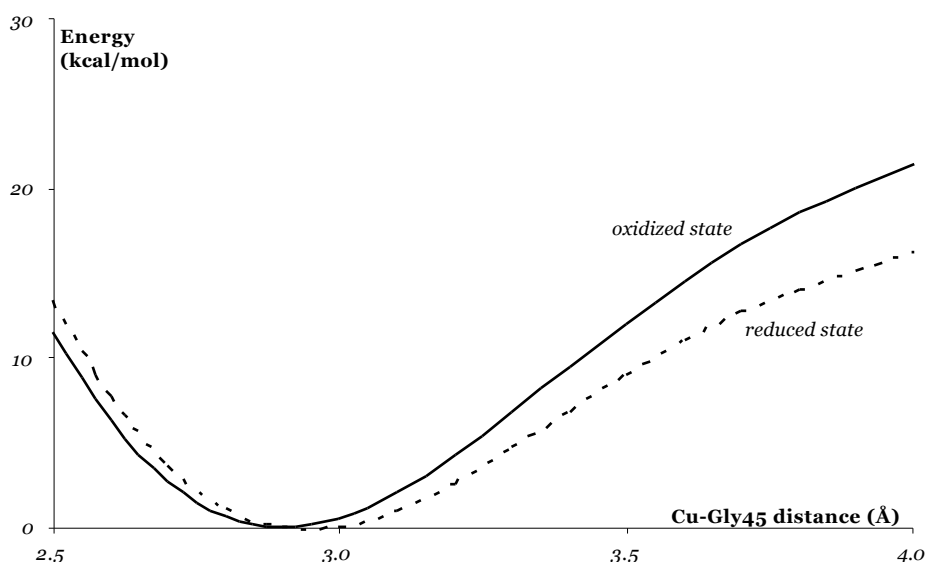


FIGURE 6.3.2. ENERGY PROFILE FOR GLY45-CU DISTANCE

<sup>a</sup> The HOMO-2 means the second orbital below the Highest Occupied Molecular Orbital (HOMO)

The energies for the reduced and oxidized state as function of the Gly45-Cu distance are given in Figure 6.3.2. The energy is given for the reduced and oxidized state relative to the lowest energy of that state, which is found in both cases at a Cu-Gly45 distance of 2.9 Å. The energy needed to put the Gly45 residue at 2.5 Å is 13.4 and 11.6 kcal/mol for the reduced and oxidized state respectively, while it costs 16.3 (*reduced state*) and 21.6 (*oxidized state*) kcal/mol to increase the Cu-Gly45 distance from 2.9 to 4.0 Å.

#### *Variation of the Met121-Cu distance*

For the reduced state with Met121 at 2.5 Å, the doubly occupied HOMO of the reduced state consists of contributions from S<sub>112</sub> (56 %) and copper (31 %). The sulphur of Met121 is not involved in the HOMO, but does have some contribution to the HOMO-1 (5 %), which consists further of contributions from copper (41 %) and S<sub>112</sub> (37 %). As expected, no such mixing in of S<sub>121</sub>-orbitals is observed when Met121 is placed at 4.0 Å from copper; only contributions to orbitals located entirely on Met121 are found.

In the oxidized state with Met121 at 2.5 Å, the singly occupied orbital consists of contributions from Cys112 (37 %), copper (31 %) and Met121 (16 %). Also in other molecular orbitals considerable contributions from Met121 are found. With Met121 at 4.0 Å, the singly occupied HOMO consists mainly of contributions of S<sub>112</sub> (51 %) and copper (27 %) with some smaller contributions (~2 %) from the neighboring nitrogens of His46 and His117.

The energies for the reduced and oxidized state as function of the Met121-Cu distance are given in Figure 6.3.3. The energy is given for the reduced and oxidized state relative to the lowest energy of that state, which is found at a Cu-Met121 distance of 3.6 Å for the reduced state and at 3.5 Å for the oxidized state. The energy needed to put Met121 at 2.5 Å is much larger than for Gly45: 22.3 kcal/mol for the reduced state and 29.8 kcal/mol for the oxidized state.

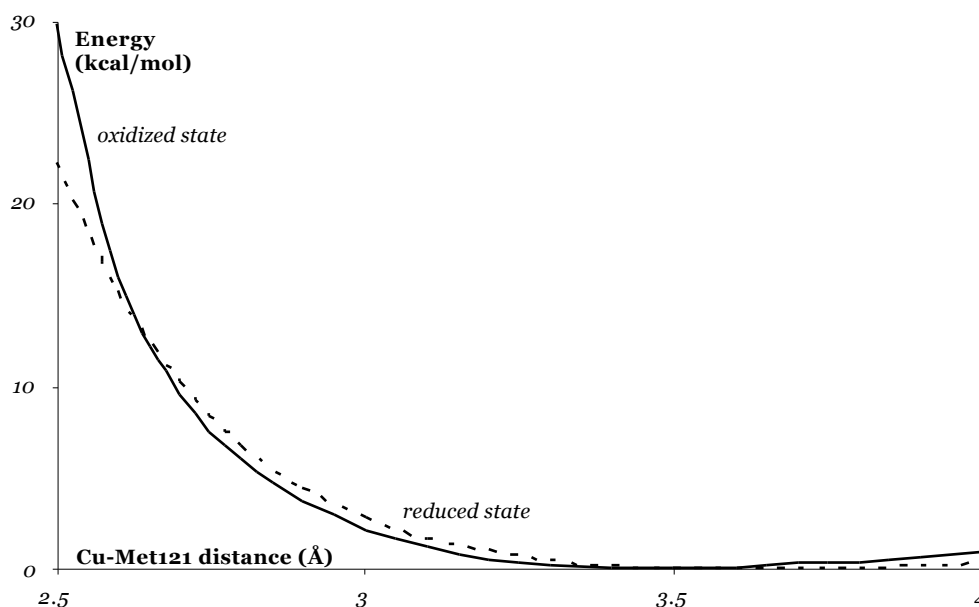


FIGURE 6.3.3. ENERGY PROFILE FOR MET121-CU DISTANCE

Around the lowest energy distance, a rather weak bonding profile is found; for instance in the oxidized state, it requires only 1.2 kcal/mol to decrease the Cu-Met121 distance from 3.5 to

3.1 Å, and only 0.9 kcal/mol to increase it to 4.0 Å. Therefore, a plateau is observed where the Met121 can move for almost an Angstrom with an energy change of approximately only 1.0 kcal/mol.

The observed plateau makes it difficult to get an accurate geometry for the active site by treating it *in vacuo*, as the influence of the protein on the Met121 residue is probably of the same order of magnitude as the energy difference at this plateau, or even higher. The inclusion of the protein in QM/MM calculations (see Section 2.3 and Chapter 9) is therefore needed to obtain correct geometries.

#### *Lowest energy orbital compositions*

In the reduced state, both with Gly45 placed at 2.9 Å from copper and with Met121 placed at 3.6 Å, the doubly occupied HOMO consists for 51 % of S<sub>112</sub> contributions and 36 % of copper contributions. At these distances, no significant mixing between copper and oxygen orbitals is found, while some small mixing is observed between Met121 and copper orbitals.

In the oxidized state for the singly occupied HOMO, the contributions of S<sub>112</sub> and copper are respectively 52 % and 27 %. Some small mixing is observed between both copper and oxygen orbitals, and copper and sulphur (Met121) orbitals.

#### *Force field parameterization*

Mixing of copper and residue orbitals is indicative of covalent interactions between copper and the residues. At equilibrium distance, no mixing of Gly45 and copper orbitals is observed in the reduced state, while there is some mixing in the oxidized state; Met121 and copper orbitals mix both in the reduced and oxidized state. Therefore, it seems there are covalent interactions between copper and Met121 in both states, between copper and Gly45 only in the oxidized state. This is reflected in the “harmonic Frost” force constants (Section 6.2), which show the same pattern; therefore, modeling the interactions in the active site of azurin with five copper-residue bonds seems justified.

## **Conclusions**

The influence of the distance of the axial groups from the copper is investigated by looking at the potential energy surface obtained by monitoring the energy as each one of them is moved away from the copper in steps of 0.1 Å. At the same time, the amount of mixing of the orbitals of the copper and the axial ligands is monitored.

The energy surface for the Gly45 residue shows a well defined minimum around 2.9 Å, both for the reduced and oxidized state, which is very close to where it is found in the crystal structures of azurin. On the contrary for the Met121 residue a very flat energy surface is found, with a change in energy of around 1 kcal/mol if the residue is moved from 3.0 to 4.0 Å. The “optimal” copper-Met121 distance is then for both the reduced and oxidized state around 3.5 Å. This enhanced freedom to move at almost no cost may presumably be the result of the absence of the protein matrix.

The amount of mixing between copper orbitals and orbitals of the oxygen of Gly45 is small, although some mixing is present when the latter is at a distance of 2.5 Å from copper. Some mixing is observed between the copper orbitals and the orbitals of the sulphur of Met121, both in the reduced and oxidized state.

## 6.4 *Vibrational frequencies*

*Validating the copper force field<sup>a</sup>*

In Section 6.2, the computed Hessians from Density Functional Theory calculations<sup>1</sup> were used to extract force constants for bonds of copper to the residues in the active site. Those Hessians provide even more information; they can be used to obtain vibrational frequencies with corresponding intensities.<sup>1</sup> For this, the Hessian  $H$  has to be transformed into mass-weighted coordinates:

$$H_{ij}^{mass\text{-}weighted} = \frac{H_{ij}}{\sqrt{M_i M_j}} \quad (1)$$

with  $M_i$  the mass of atom  $i$ . After diagonalization of the mass-weighted Hessian, the eigenvalues  $\bar{\omega}$  can be transformed quite easily into the frequencies  $\omega_i$  with corresponding normal mode vectors:

$$\omega_i = \frac{1}{2\bar{\omega}} \sqrt{\bar{\omega}_i} \quad (2)$$

This procedure works for any Hessian in Cartesian coordinates. Therefore, it can be used directly to check the vibrational frequencies that would be the result of the interactions of the five bonds (see Section 6.2).

The vibrational frequencies obtained in that way would refer to an isolated system containing only those five bonds. However, they will be applied to the active site in a protein and one wants to compare them with experimentally observed frequencies, which are usually obtained at finite (room) temperature. For a meaningful comparison, one therefore should perform Molecular Dynamics simulations<sup>160</sup> at that temperature and extract the values for the frequencies from those simulations. In principle, the same procedure could be used as described above, i.e. calculate the Hessian for all coordinates, transform it to the mass-weighted coordinate system and diagonalize it. However, this would be an enormous task; the Hessian matrix should be averaged over time, while the time needed for just one Hessian would be enormous due to the system size of several thousands of atoms. A more efficient way of extracting the values for the vibrational frequencies can be obtained by monitoring the copper-residue bond lengths in time. Transforming the values from real space to frequency space by taking the Fourier transform results then directly in the values for the vibrational frequencies resulting from those bonds. As these values are obtained in simulations at a particular temperature, they can be compared directly with the experimentally observed values.

---

<sup>a</sup> Joint study by M. van den Bosch (MD simulations) and M. Swart (DFT calculations)

## Computational details

The Density Functional Theory calculations were performed with the ADF program (versions 2.3.3, 1999.03, 2000.02)<sup>177,178,187</sup> on a cluster of either IBM RS/6000 or Pentium-Linux boxes. In all calculations the Becke<sup>120</sup>-Perdew<sup>121</sup> exchange-correlation potential was used in a triple zeta basis set with polarization functions (TZP). The Hessian matrix (a matrix containing the second derivatives of the energy with respect to atomic coordinates) was determined for copper and its neighboring atoms in the five residues in the active site. In total the Hessian matrix has therefore a dimension of  $18 \times 18$ , and leads to 18 vibrational frequencies. For a molecule consisting of 6 atoms (and therefore also to 18 vibrations), six of these would correspond to translation or rotation of the molecule and have a frequency of zero; however, in this case we are dealing with a sub-part of the Hessian of the molecule (which has a size of  $183 \times 183$  for the active site consisting of 61 atoms). I.e., the system of the six atoms is connected to the residues and therefore does not have the six modes with frequency zero.

The coordinates for the azurin variants were taken from PDB-files that were obtained from the Brookhaven Protein Database {F114A *Pseudomonas aeruginosa* (1azn), N47D *Pseudomonas aeruginosa* (1azr), wildtype *Pseudomonas aeruginosa* pH5.5 (4azu; hereafter referred to as *wt1*), wildtype *Pseudomonas aeruginosa* pH9.0 (5azu; *wt2*)} or from personal communication<sup>280</sup> {*Alcaligenes denitrificans* M121Q AD<sup>42</sup>, N47L<sup>84</sup> AD}.

The Molecular Dynamics simulations and analyses were performed with the Gromacs suite<sup>279</sup> (version 2.1) on a SGI R12000 and a DEC/Alpha-linux cluster. The Multipole Derived Charges<sup>183</sup> were used for the active site atoms and the IntraFF force constants for *wildtype Pa* azurin (see Section 6.2) for bonds of copper to the five residues in the active site. For all other force field parameters, the standard GROMOS96<sup>127</sup> values were used, where the high Lennard-Jones<sup>160</sup> repulsion parameter for Cu<sup>127</sup> was used throughout. The simulations consisted of 1 nanosecond at a temperature of 300K and a pressure of 1 bar, which were kept constant using the Berendsen<sup>166</sup> thermo-barostat. The protein was solvated in 4152 (SPC/E<sup>281</sup>) water molecules. The cut-off radius was set at 14 Å, used for Lennard-Jones and electrostatic interactions, and 8 Å for the neighborlist that is updated every ten time steps. No constraints were put on bonds, angles or dihedrals in the active site, while the bond lengths were constrained for the rest of the protein, and the bond lengths and angle for the solvent molecules using the LINCS procedure.

## Results

The vibrational frequencies from the DFT Hessians are given in Table 6.4.1 for both the reduced and oxidized state. Given also are the intensities for the frequencies, which are obtained from taking the derivative of the dipole moment with respect to the atomic coordinates. The frequencies for the reduced states of the six azurin variants are mostly lower than those found for the oxidized state. This is in agreement with the force constants for the copper-residue bonds, which were extracted from these Hessians and found to be smaller in the reduced than in the oxidized state. For all but the F114A-mutant, an intense frequency around 400 cm<sup>-1</sup> is observed (intensity around 30 km/mol), while even more intense frequencies are observed in the 900-1100 cm<sup>-1</sup> region. Contributions of the Cu-Cys

bond to the normal mode vectors<sup>a</sup> are mainly found in the 280-370 cm<sup>-1</sup> region, with some smaller contributions for the frequencies in the region of 400 to 500 cm<sup>-1</sup>.

TABLE 6.4.1. FREQUENCIES (CM<sup>-1</sup>) AND *INTENSITIES* (KM/MOL) FROM DFT

<i>wt1</i>		<i>wt2</i>		<i>n47d</i>		<i>f114a</i>		<i>n47l</i>		<i>m121q</i>	
<i>reduced state</i>											
84	1.8	86	2.9	67	2.3	56	4.4	82	2.0	93	2.6
*109	9.7	*117	6.9	*89	9.8	70	3.8	84	8.5	103	7.6
*124	1.2	*128	3.5	*119	1.4	*119	1.3	*157	0.5	*126	0.9
*158	0.3	*157	0.6	190	1.3	142	11.1	169	1.2	*189	1.7
196	1.3	180	1.3	*195	0.3	189	1.0	*173	0.7	211	6.9
218	7.2	223	7.4	229	6.4	293	7.6	238	5.6	*240	6.5
300	12.4	292	13.2	313	8.7	326	19.4	369	6.7	380	4.7
406	4.2	417	33.0	412	30.5	383	1.4	409	33.4	*400	27.2
417	25.0	*425	2.3	448	3.5	*390	11.8	420	3.5	*405	2.4
425	9.6	*436	14.8	450	13.3	411	1.8	452	12.7	417	27.0
*435	14.3	460	0.2	454	0.2	414	5.3	462	0.5	435	8.8
453	0.2	471	3.9	481	0.2	417	2.3	491	0.5	445	10.6
476	0.3	491	0.4	507	4.8	426	20.5	529	3.9	1071	10.2
1056	5.6	691	3.9	929	3.1	910	65.3	929	1.1	1084	8.7
1078	12.7	1067	10.0	954	4.3	914	43.3	1056	113.6	1086	61.6
1090	19.3	1075	5.8	1066	18.1	995	18.4	1078	13.1	1103	99.5
1090	89.2	1080	102.2	1070	20.4	1091	46.8	1088	2.7	1104	5.9
1097	6.3	1095	7.3	1090	95.0	1117	17.7	1101	8.7	1114	6.0
<i>oxidized state</i>											
93	3.2	93	5.8	90	6.9	*57	10.2	97	6.1	100	4.5
118	3.0	123	1.1	104	1.3	92	8.9	108	1.3	119	3.6
*151	11.4	*154	5.2	*145	12.8	*129	4.2	179	1.5	*158	5.9
*183	6.0	*181	4.9	195	0.7	136	1.4	*186	3.9	*207	5.8
203	0.8	189	2.1	*212	5.6	191	1.5	195	4.5	*209	3.4
215	7.5	220	6.3	228	5.6	288	7.8	237	5.6	*232	11.5
296	7.9	287	10.4	313	8.2	320	6.9	366	7.4	375	4.9
*408	0.1	416	30.6	411	39.1	368	1.4	411	32.4	*406	31.2
*418	31.2	*418	1.9	433	0.4	387	5.0	424	2.3	*408	6.2
*420	6.3	431	0.9	448	1.5	409	4.2	444	2.3	417	23.2
428	1.3	458	2.1	454	2.6	414	0.9	460	0.3	436	2.1
451	4.0	485	4.4	480	2.5	416	1.7	490	2.8	442	5.9
476	2.9	516	13.9	512	2.5	424	25.7	531	7.2	1085	14.1
1080	13.8	711	9.4	945	22.1	912	51.4	958	12.3	1096	53.1
1090	85.7	1080	56.1	985	4.9	918	72.9	1059	137.8	1101	97.8
1093	76.5	1082	79.6	1076	17.0	1003	22.2	1093	8.0	1105	119.1
1106	0.4	1096	25.3	1088	36.3	1099	39.7	1108	5.2	1120	7.7
1109	16.7	1104	2.6	1099	143.2	1129	21.4	1113	16.5	1138	14.0

**Bold** frequencies have more than 40 % contributions of the Cu-Cys(112) bond<sup>a</sup>

Frequencies indicated with an asterisk (\*) have contributions between 10 and 40 %

<sup>a</sup> Obtained by taking the overlap of the normal mode vector and the Cu-Cys112 bond vector

The frequencies as obtained from the FF Hessian matrices (coming from the five copper-residue bonds only; see Section 6.2) are given in Table 6.4.2; in this case, only frequencies can be obtained. The frequencies from the five classical bonds are observed in the low-frequency region (100-570  $\text{cm}^{-1}$ ), both for the reduced and the oxidized state. As the force constants have a higher value in the oxidized state, the frequencies are found at higher values than in the reduced state. The contribution of the Cu-Cys bond to the normal mode vectors is found mainly in the 270-420  $\text{cm}^{-1}$  region.

TABLE 6.4.2. FREQUENCIES ( $\text{CM}^{-1}$ ) FROM THE CLASSICAL COPPER-RESIDUE BONDS

<i>wt1</i>	<i>wt2</i>	<i>n47d</i>	<i>f114a</i>	<i>n47l</i>	<i>m121q</i>
<i>reduced state</i>					
121	121	101	107	128	170
162	158	143	126	195	205
*246	<b>249</b>	*244	146	<b>355</b>	288
<b>296</b>	<b>291</b>	<b>332</b>	*258	422	<b>391</b>
*363	<b>371</b>	*508	<b>329</b>	<b>447</b>	<b>515</b>
<i>oxidized state</i>					
132	132	128	114	149	206
175	172	179	131	217	249
<b>287</b>	<b>286</b>	<b>313</b>	152	<b>390</b>	*349
<b>357</b>	<b>356</b>	<b>366</b>	<b>279</b>	493	<b>415</b>
*432	*438	*564	<b>339</b>	<b>504</b>	<b>569</b>

**Bold** frequencies have more than 40 % contributions of the Cu-Cys(112) bond

Frequencies indicated with an asterisk (\*) have contributions between 10 and 40 %

The Molecular Dynamics simulations were performed for *wt1* azurin only in either the reduced or oxidized state. In these simulations, both the active site and the protein are stable (see Figures 6.4.1 and 6.4.2).

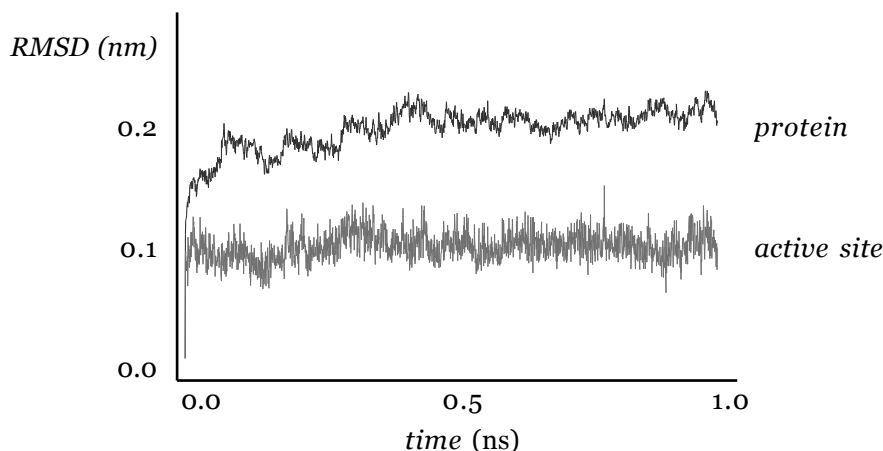


FIGURE 6.4.1. RMSD OF REDUCED STATE DURING MD SIMULATION

In these figures, the root mean square displacement (RMSD) of the atoms from their initial positions is given as function of the simulation time, either for the active site or the protein.



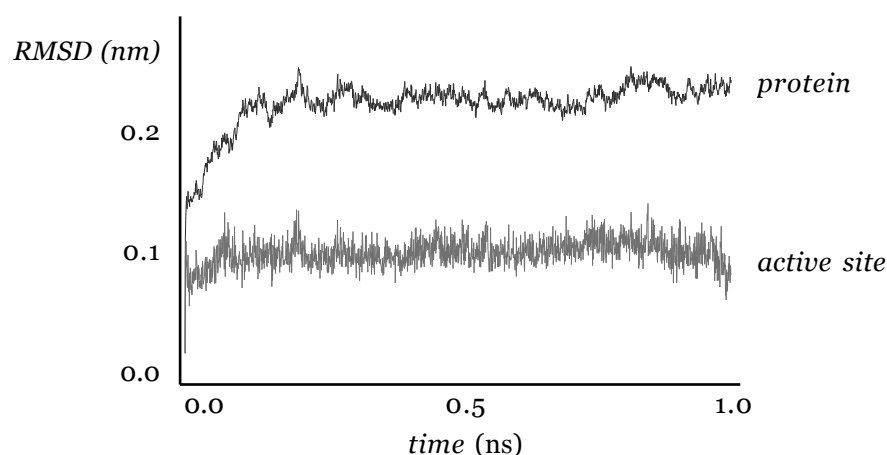


FIGURE 6.4.2. RMSD OF OXIDIZED STATE DURING MD SIMULATION

The RMSD of the active site atoms is in both the reduced and the oxidized state just below 1 Å, while it is around 2 Å for the C-alpha atoms of the complete protein. The stability of the site is confirmed by the average copper-residue distances as obtained in the MD simulation over a period of 500 ps (see Table 6.4.3). The difference between the experimental and simulated distances is small (less than 0.1 Å) and at most of the same magnitude as the experimental uncertainty.

TABLE 6.4.3. AVERAGE COPPER-RESIDUE DISTANCES (Å) FROM MD SIMULATIONS

	<i>reduced state</i>		<i>oxidized state</i>	
	<i>X-ray</i>	<i>MD</i>	<i>X-ray</i>	<i>MD</i>
R(Cu-Gly45)	$3.02 \pm 0.07$	3.04	$2.97 \pm 0.09$	3.05
R(Cu-His46)	$2.13 \pm 0.09$	2.09	$2.08 \pm 0.06$	2.08
R(Cu-Cys112)	$2.29 \pm 0.01$	2.31	$2.24 \pm 0.04$	2.29
R(Cu-His117)	$2.09 \pm 0.09$	2.07	$2.01 \pm 0.07$	2.07
R(Cu-Met121)	$3.25 \pm 0.07$	3.34	$3.15 \pm 0.07$	3.29
R(Cu-N <sub>2</sub> S plane)	0.09	0.12	0.08	0.08

The vibrational frequencies for the Cu-Cys and both Cu-His bonds as obtained in the MD simulations<sup>a</sup>, are given in Figure 6.4.3. The vibrational frequencies are all found below the 450 cm<sup>-1</sup>, where the Cu-His frequencies are found mainly in the 200-350 cm<sup>-1</sup> region and the Cu-Cys frequencies around 280 cm<sup>-1</sup> and mainly from 350-400 cm<sup>-1</sup>. They compare very well with the frequencies observed in Resonance Raman spectra of *wildtype* azurin (see Figure 6.4.3).

<sup>a</sup> from smoothed Fourier transform

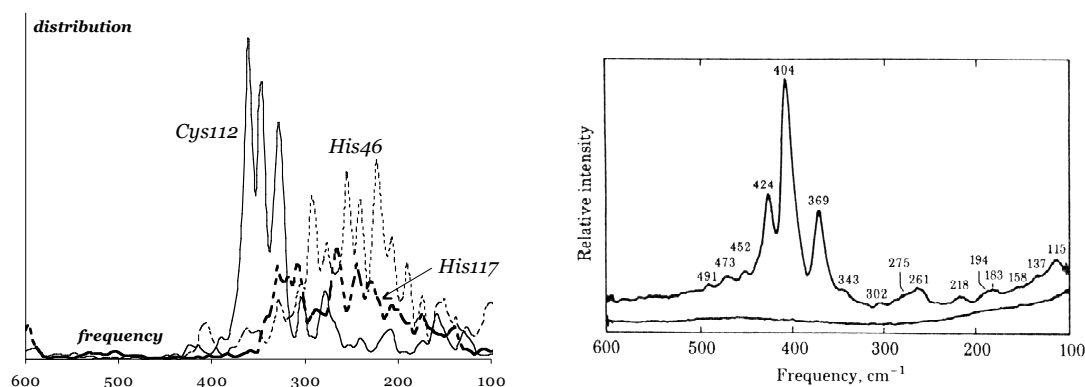


FIGURE 6.4.3. VIBRATIONAL FREQUENCIES FOR CU-RESIDUE BONDS FROM MD SIMULATIONS (LEFT) AND RESONANCE RAMAN SPECTRUM OF WILDTYPE AZURIN (TAKEN FROM REF 62)

## Discussion

The frequencies obtained in Density Functional Theory<sup>1</sup> calculations are all found below  $1200\text{ cm}^{-1}$  (see Table 6.4.1), with the most intense ones observed around  $400\text{ cm}^{-1}$  and above  $900\text{ cm}^{-1}$ . The normal mode vectors for the frequencies with higher values ( $> 900\text{ cm}^{-1}$ ) consist exclusively of contributions from copper, nitrogen and oxygen atoms. Some preliminary tests where classical N-Cu-O angles have been added to the force field seem to indicate that they may be partly responsible for these higher frequencies.

The highest frequency where a sulphur atom is contributing to the normal mode vector is observed at  $516\text{ cm}^{-1}$  (*wt2 in oxidized state*; Met121 sulphur). The largest contributions to a normal mode vector by the Cu-Cys bond are observed primarily in the region of  $280\text{--}370\text{ cm}^{-1}$ , with some smaller contributions observed for frequencies between  $400\text{--}500\text{ cm}^{-1}$ . Depending on the copper-sulphur distance, these frequencies shift up- or downwards. For instance, for *wt1* and *wt2* azurin with a Cu-S distance of  $\sim 2.27\text{ \AA}$ , the seventh frequency is found around  $290\text{ cm}^{-1}$ , while for the N47D and F114A mutants (with a smaller Cu-S distance of  $2.22\text{ \AA}$ ) it is increased to around  $320\text{ cm}^{-1}$  and for the N47L and M121Q mutants with an even smaller distance of  $2.13\text{ \AA}$  the frequency occurs around  $370\text{ cm}^{-1}$ .

A similar but less pronounced trend is observed for the average of the Cu-His distances in a certain azurin variant and the thirteenth frequency. For the N47L mutant with an average Cu-N distance of  $1.96\text{ \AA}$  it is observed at  $531\text{ cm}^{-1}$ , while for the F114A mutant with an average of  $2.29\text{ \AA}$  it is found at only  $424\text{ cm}^{-1}$ . It is found inbetween these two extremes for *wt1*, *wt2* and the N47D mutant. For M121Q-azurin, it is found at  $1085\text{ cm}^{-1}$  due to mixing in of the Cu-Gln121 bond, which has a distance of only  $2.25\text{ \AA}$ .

The frequencies from the five copper-residue bonds alone are all lower than  $570\text{ cm}^{-1}$  (see Table 6.4.2). The positioning of these frequencies depends to a large extent on the Cu-residue distances, as can be seen from the frequencies observed in either the N47D or the F114A mutant. In the former, the in-plane copper-ligand distances are relatively small leading to higher values for the observed frequencies. The latter, with rather large copper-His distances, exhibits relatively low frequencies. In all cases however, the highest frequency has a considerable contribution from either one of the Cu-His bonds, with an usually smaller contribution from the Cu-Cys bond. The second highest observed frequency consists mainly of contributions from the Cu-Cys bond and Cu-His bonds, while the third frequency (around  $290\text{ cm}^{-1}$ ) consists mainly of contributions from the Cu-Cys bond. The two lowest

frequencies consist mainly of contributions from the axial residues, with some small contributions from either the Cu-Cys or the Cu-His bonds.

In the Molecular Dynamics simulations of 1 ns, both the active site and the protein remain stable after an initial equilibration period of some 500 ps (reduced state) or 200 ps (oxidized state). In both cases, the active site reaches stabilization faster than the protein, where a smaller root mean square deviation is observed for the oxidized state than for the reduced state. As the protein structure was determined experimentally for the former, this does not come as a surprise. The copper-residue distances observed during the second half of the 1 ns simulations are found close to the ones found in the X-ray data. For the in-plane ligands, a good agreement is found between simulated and experimental distances. For the axial residues the differences are somewhat larger, but remain within the experimental accuracy of 0.1-0.2 Å of the X-ray data.

The distance of the copper from the N<sub>2</sub>S plane shows a good agreement between the simulated and experimental values. In both experiment and simulation, the copper remains above the plane at some 0.1 Å. As this distance has not been modeled in any kind of bonding interactions, it may serve as valuable check for determining the validity of the copper force field. The good agreement between experimental and simulated values indicates that the copper force field described in Sections 6.1 and 6.2 gives a proper description for the bonding interactions in the active site of azurin.

This is confirmed by the vibrational frequencies that are obtained in the MD simulations. For the Cu-Cys bond, they are observed mainly in the 300-400 cm<sup>-1</sup> region with some smaller ones observed around 280 cm<sup>-1</sup>, while for the Cu-His bonds they are observed at (mainly) lower or higher frequencies (see Figure 6.4.3). Although this figure doesn't provide intensities, as there is a large resemblance between the frequencies and normal modes observed in DFT calculations, the copper force field and the MD simulations, it seems justified to use the intensities obtained in the DFT calculations and apply them to the frequencies in a particular frequency range. Therefore, the intense frequencies observed in Resonance Raman studies around 400 cm<sup>-1</sup>, which were attributed to copper-sulphur interactions, show up also in either DFT, copper force field or MD simulations results.

## Conclusions

This section handles the vibrational frequencies that are obtained from either Density Functional Theory (DFT) calculations, copper force field calculations where the force field parameters have been extracted from Hessian matrices obtained in the DFT calculations, or from Molecular Dynamics simulations. A general agreement is observed for the frequencies resulting from the copper-residue bonds, which are all found below 1200 cm<sup>-1</sup>.

In the DFT calculations, the most intense frequencies are observed around 400 cm<sup>-1</sup> and above 900 cm<sup>-1</sup>. Generally speaking, the frequencies are observed at higher values in the oxidized state than in the reduced state. The frequencies around 400 cm<sup>-1</sup> can be attributed entirely to contributions from either the Cu-Cys or the Cu-His bonds.

Force field parameters for the copper-residue bonds have been extracted from the DFT Hessian matrices, which give frequencies between 100-600 cm<sup>-1</sup>. The values of these frequencies are shown to depend critically on the in-plane copper-ligand distances, just like was found for the force constants for these in-plane bonds. A decreasing copper-residue

distance leads to higher values for the force constant, and therefore to higher values for the frequencies resulting from the copper force field.

As final check, Molecular Dynamics simulations of 1 ns were performed to check the validity of the copper force field. In these simulations, both the active site and the protein remain stable with root mean square deviations (RMSD) of around 1.0 and 2.0 Å respectively. This stability is confirmed by the copper-residue distances, which show a good agreement with the ones observed in the X-ray data. The distance of the copper from the N<sub>2</sub>S plane also showed agreement between simulated and experimental values. This may be a valuable check for the validity of the copper force field, as this distance was not included in any of the bonding interactions of the copper force field.

The vibrational frequencies for the in-plane ligands were obtained in the simulations. The frequencies from the Cu-Cys bond were found primarily between 300 and 400 cm<sup>-1</sup>, while those for the Cu-His bonds were found either at lower or higher values. Therefore, a very good agreement is observed between the vibrational frequencies observed either by Density Functional Theory, copper force field or Molecular Dynamics simulations.